Geometry of the Phosphate Group and Its Interactions with Metal Cations in Crystals and *ab Initio* Calculations

Bohdan Schneider,* Martin Kabeláč, and Pavel Hobza

Contribution from the J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, 18223 Prague, Czech Republic

Received June 24, 1996. Revised Manuscript Received September 30, 1996[⊗]

Abstract: The phosphate geometry was studied in crystal environment by analyzing 178 crystal structures and *in vacuo* by *ab initio* calculations of (di)hydrogen and dimethyl phosphates and a diphosphate model system, $CH_3OP(O_2)O(CH_2)_3OP(O_2)OCH_3$ (diPh), at the MP2 and HF levels. Since charge determines the phosphate stereochemistry, uncharged, −1, and −2 phosphates were analyzed separately. The P═O and O═P═O bonding parameters depend only on charge and the number of carbon substituents while the remaining parameters depend also on substituent type; the largest sensitivity was observed for the C−O and C−O−P parameters. An acceptable agreement between the crystal and *ab initio* geometries was obtained only when basis sets contained polarization functions and a model system included a counterion (Na^+). In uncharged and −1 phosphates, the C−O−P−O(−C) torsion angles prefer the ±*sc* regions. These phosphates substituted by cyclic C_{sp^3}, noncyclic C_{sp^3}, or C_{ar} carbons have characteristic torsion distributions. Conclusion from both analysis of crystal data and from theoretical calculations is that the internal phosphate geometry is sensitive to a counterion position as well as to values of the C−O−P−O(−C) torsion angles. Most metal cations interact directly with the charged oxygens in −1 phosphates while in many −2 phosphates, this interaction is mediated by water molecules. The distributions of Na^+ around −1 and −2 phosphates are localized into two principal sites which lie outside the O═P═O plane and interact with only one of the charged oxygens. In contrast, theoretically predicted positions are located symmetrically between the charged oxygens in the O═P═O plane. The cation positions around dimethyl phosphate and a model diPh compound mimicking the B-DNA backbone were virtually identical. The discrepancy between these theoretical positions and the sites derived from the crystal data was significantly reduced by incorporation of a single water molecule into the theoretical model. Therefore, both cations and polar particles, particularly water molecules, should be considered when properties of the phosphate group are described.

Introduction

The phosphate group is an essential component in a variety of biological macromolecules, including its predominant role in nucleic acids and lipids as well its importance in coenzymes and secondary messengers. The phosphodiester linkage of nucleic acids has been recently surveyed using high-resolution nucleoside and nucleotide crystal structures[1] as well as lower resolution structures of oligonucleotides[2] with the aim of reevaluating the bond distances and angles used as geometric standards in crystallographic refinement and to give an overview of observed double helical conformations. We were able to complement these studies through a comprehensive survey of a broader class of phosphate structures and analyze the relationship among their geometries, charge, type of chemical substitution, and interacting counterion type and position. Moreover, for the first time, we systematically compare the averaged crystal geometries to reliable nonempirical theoretical models.

The smallest molecule which can adequately describe some conformational features of the DNA backbone and of the charged head of phospholipids is dimethyl phosphate (DMP). Several theoretical studies of the DMP^- and of similar compounds were primarily performed at the Hartree−Fock level (HF) using small- or medium-sized atomic orbital basis sets.[3−21] However, higher level calculations considering the correlation energy should be performed because the correlation energy could significantly affect the geometrical characteristics of the phosphate group. Besides the level of theoretical description, another limitation of previous studies is the size of the most commonly used model dimethyl phosphate. Even when it contains both

* To whom correspondence should be addressed. E-mail: schneider@jh-inst.cas.cz. Fax: +42-2-858-2307.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1996.

(1) Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S.-H.; Olson, W. K.; Berman, H. M. *J. Am. Chem. Soc.* **1996**, *118*, 518−529.
(2) Schneider, B.; Neidle, S.; Berman H. M. *Biopolymers,* in press.
(3) Newton, M. D. *J. Am. Chem. Soc.* **1973**, *95*, 256−258.
(4) (a) Guan, Y; Choy, G. S.-C.; Glaser, R.; Thomas, G. J., Jr. *J. Phys. Chem.* **1995**, *99*, 12054−12062. (b) Guan, Y.; Wurrey, C. J.; Thomas, G. J., Jr. *Biophys. J.* **1994**, *66*, 225−235. (c) Guan, Y.; Thomas, G. J. Jr. *J. Mol. Struct.* **1996**, *379*, 31−41.
(5) Deerfield, D. W., II; Pedersen, L. G. *J. Biomol. Struct. Dyn.* **1995**, *13*, 167−180.
(6) Florián, J.; Baumruk, V.; Štrajbl, M; Bednárová, L.; Štěpánek, J. *J. Phys. Chem.* **1996**, *100*, 1559−1568.
(7) Deleted on revision.
(8) Perahia, D.; Pullman, B. *Biochem. Biophys. Acta* **1976**, *435*, 282.
(9) Gorenstein, D. G.; Kar, D.; Luxon, B. A. *J. Am. Chem. Soc.* **1976**, *98*, 1668−1673.
(10) Gorenstein, D. G.; Kar, D. *J. Am. Chem. Soc.* **1977**, 99, 672−676.
(11) Clementi, E.; Gorongui, G., Lelj, F. *J. Chem. Phys.* **1979**, *70*, 3726−3729.
(12) Nanda, R. K.; Govil, G. *Theor. Chim. Acta* **1975**, *38*, 71−77.
(13) Pullman, B.; Gresh, N.; Berthod, H. *Theor. Chim. Acta* **1975**, *40*, 71−74.
(14) Marynick, D. S.; Schaeffer, H. F., III. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3794−3799.
(15) Berthod H.; Pullman, A. *Chem. Phys. Lett.* **1977**, *46*, 249−252.
(16) Pullman, A.; Pullman B. *Theor. Chim. Acta* **1978**, *47*, 175−191.
(17) Liebmann P.; Loew, G.; McLean, A. D.; Pack, R. G. *J. Am. Chem. Soc.* **1982**, *104*, 691−697.
(18) Landin, J.; Pascher, I. *J. Phys. Chem.* **1995**, *99*, 4471−4485.
(19) Hadzi, D.; Hodosczek, M.; Grdadolnik, J.; Avbelj, F. *J. Mol. Struct.* **1992**, *266*, 9−19.
(20) Liang, C.; Ewig, C. S.; Stouch, T. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1993**, *115*, 1537−1545.
(21) Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1995**, *99*, 11591−11599.

S0002-7863(96)02115-4 CCC: $12.00 © 1996 American Chemical Society

phosphodiester bonds present in the DNA backbone, it is too simple to model DNA conformational flexibility, repulsion of nearby phosphate groups, and interactions with cations so that a larger (and computationally more demanding) model is studied here.

Due to the charge of the phosphate group its interactions with cations are energetically important. The metal cations influence the biological function of double helical DNA[22] and change its melting curves.[23,24] Without closely attached cations, unshielded interaction of negatively charged phosphates would be strongly repulsive, drastically decreasing the stability of the double helical DNA. The phosphate group in biologically relevant molecules is charged under physiological conditions, and most of its charge is compensated for by light metal cations such as Na^+ or Mg^{2+}. The level of organization of these cations around DNA is, however, unknown. Since oligonucleotide crystal structures are solved at a relatively low crystallographic resolution, metal cations are only rarely located. It is not, therefore, possible to decide whether they are ordered or randomly scattered around the DNA helical backbone. As a result, oligonucleotide structures can neither confirm nor disprove some of the polyelectrolyte theories.[25,26] It should be stressed that if we demonstrate that metal cations concentrate into well-localized sites and therefore some conclusions of these polyelectrolyte theories can be tested, it will help to interpret the behavior of the DNA double helix in solution, as well as to decipher results of computer simulations.

In contrast to oligonucleotide crystals, even the light cations are ordered and therefore visible in crystals of organic phosphates with low molecular weight. A pioneering paper by Alexander *et al.*[27] showed that metal cations are partially ordered around the anionic phosphinyl group $(-PO_2-)^-$. A limited number of crystal structures available at that time did not allow the study of distribution of a specific cation around the whole phosphate group. An increase in available structures allowed us to focus on the distribution of Na^+ and study phosphates with charges -1 and -2 separately. This analysis complemented by results of extensive *ab initio* calculations revealed the nature of the behavior of Na^+ around the phosphate group.

This work presents a first systematic examination of geometries of the phosphate group in crystal environment as well as *in vacuo*. More than 170 phosphate crystal structures were analyzed as a function of charge, substitution, and cation position and are compared to geometries of three phosphate-containing model systems calculated by *ab initio* quantum chemical methods performed at the Hartree−Fock and correlated levels. Special attention was paid to analysis of distributions of metal cations around the phosphate group in both crystals and *ab initio* model systems.

Methods

1. Analysis of crystal structures. 1.1. Selection of Structures. The Cambridge Structural Database (CSD)[28] was searched for crystal structures containing the phosphate group using the program QUEST.[29] The database search was tailored to retrieve phosphate fragments with charges of -2, -1, or 0, and substituted by either hydrogen(s) or carbon(s) (Figure 1 a). Only structures with crystallographic *R*-factors better than 6% and with average estimated standard deviations (esd) of C−C bond lengths no greater than 0.01 Å were used in order to minimize effects of the experimental errors on the analyzed parameters. These values represent a compromise between the accuracy and number of available structures. The program GSTAT[29] was used to calculate bond lengths and valence as well as torsion angles of released compounds.

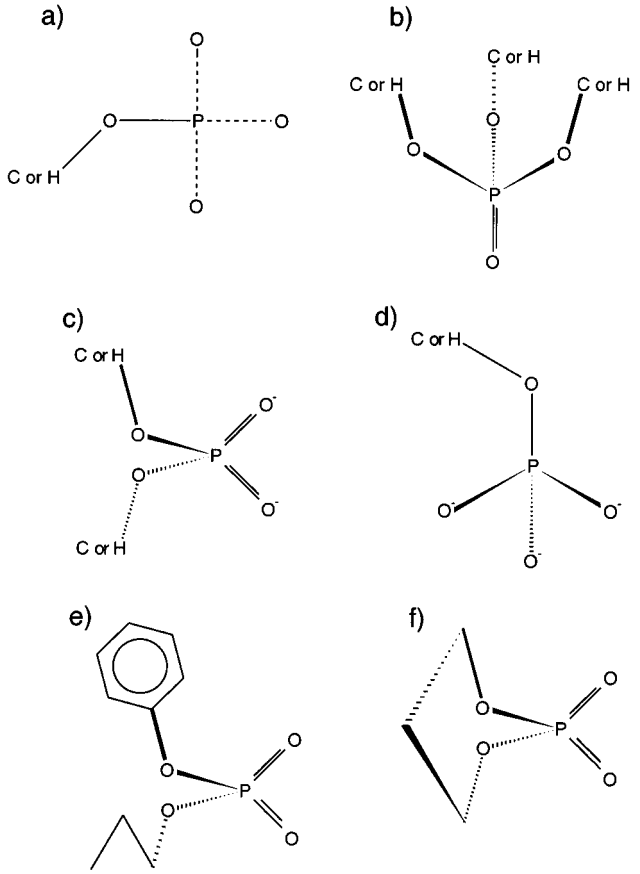

Figure 1. Fragment searched in the Cambridge Crystallographic Database (a) and structural formulas of the uncharged (b), -1 (c), and -2 (d) phosphates. The formula in (e) shows a -1 phosphate with acyclic aliphatic and aromatic substituents and the formula in (f) a cyclic phosphate.

1.2. The Phosphate Geometry. In this study, we considered all P−O bond distances and C−O distances in the immediate vicinity of the phosphorus, all bond angles involving P, and torsion angles C−O−P−O. We used symbols "P═O" and "P−O" to distinguish bonds to phosphoryl oxygens (oxy-anions) from bonds to carbon- or hydrogen-substituted oxygens, respectively.

The stereochemistry of the phosphate group depends on its charge, so the phosphates were divided into three classes on the basis of their charge (Figure 1b−d). The influence of substituents on the phosphate geometry was further examined by creating smaller classes of structures substituted by chemically more similar groups. Only classes containing more than four structures were considered. Arithmetic means and esds of all parameters were calculated in each class. The means of the presumably similar classes were compared, and only classes with significantly different means were kept separate while the others were merged, forming a larger class. In a newly created class, means and esds of the geometric parameters were recalculated and compared with these values in the remaining classes to guarantee consistency of the classification.

(22) Eichhorn, G. L. *Metal Ions in Genetic Information Transfer*; Elsevier: New York, 1981.
(23) Duguid, J. G.; Bloomfield, V. A.; Benevides, J. M.; Thomas, G. J., Jr. *Biophys. J.* **1995**, *69*, 2623−2641.
(24) Duguid, J. G.; Bloomfield, V. A. *Biophys. J.* **1995**, *69*, 2642−2348.
(25) Manning, G. S. *Q. Rev. Biophys.* **1978**, *11*, 179−246.
(26) Anderson, C. F.; Record, T. R., Jr. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 423−465.
(27) Alexander, R. S.; Kanyo, Z. F.; Chirlian, L. E.; Christianson, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 933−937.
(28) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr.* **1979**, *B35*, 2331−2339.
(29) Allen, F. H.; Davies, J. E.; Galloy, J. J.; Johnson,O.; Kennard,O.; Macrae, C.F.; Mitchell, E. M.; Mitchell, G. F.; Smith J. M.; Watson, D. G. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187−204.

The equality of two means was evaluated by the *t*-test. Since the compared distributions often had significantly different esds, the *t*-test modified by Fisher and Behrens[30] was applied. The use of this modification of the *t*-test for analysis of the structural data is described in detail elsewhere.[31] Two means were said to be significantly different when the probability that they were the same was less than 0.05.

The Klyne−Prelog nomenclature[32] is used to describe torsion angles. The following abbreviations are used to describe conformational regions: +*sc*, 30−90°; +*ac*, 90−150°; *ap*, 150−210°; −*ac*, 210−270°; −*sc*, 270−330°; *sp*, 330−30°. The program STATISTICA[33] was used to calculate the unweighted arithmetic means and their esds and for histogram plots.

1.3. Definition of Standards. The mean values of the geometric parameters of the final classes of structures were compared to the means determined for specially selected classes of structures which are hereafter labeled as "standards". The standard classes were defined separately for charges 0, −1, and −2 from chemically consistent phosphate structures which all had substituents of the same type. The standard for charge 0 was formed from phosphates substituted by three aliphatic carbon substituents, each forming just one bond to the phosphate so that they are "acyclic" and do not form cyclic phosphates, with the structural formula in Figure 1f. For charge −1, the standard was formed from structures with two aliphatic acyclic carbon substituents, and for the charge of −2 from structures with one aliphatic acyclic carbon substituent. For instance, a structure was excluded from the standard class when it had one aliphatic substituent and one aromatic substituent (as in Figure 1e).

1.4. Distributions of Metal Cations around the Phosphate Group. The structures containing cations Na^+, K^+, and Mg^{2+} were divided into classes with the phosphate charges −1 and −2. Contacts closer than 3.00 Å between cations and neighboring atoms were calculated by applying all the space group operations and *x*, *y*, *z* translations[34] and categorized by type of interacting moiety such as contacts to phosphates, water, and other parts of the crystal asymmetric unit.

The number of structures containing sodium cations was sufficient to perform further analysis of cation spatial distributions. Phosphates with charges −1 and −2 were again treated as separate classes. All cation positions closer than 4.60 Å from the phosphorus atom were generated by applying the symmetry operations of a particular crystal structure. The phosphate groups with their associated sodium cations were superimposed[35] over a template dimethyl phosphate molecule using the five phosphate atoms. The analysis of sodium distributions was based on 31 independent Na^+ positions from −1 phosphates and on 22 such positions from −2 phosphates.

The $[-O-(PO_2)-O-]^-$ fragment has two pairs of interchangeable oxygens; two charged and two ester oxygens are potentially the same. Similarly, the $[-O-(PO_3)]^{2-}$ fragment has three interchangeable charged oxygens. Due to these possible permutations of the oxygen positions each $[-O-(PO_2)-O-]^-$ fragment had to be overlapped twice over the template by permuting the interchangeable oxygen atoms of a fragment with the oxygens of the template. The resulting Na^+ distribution around −1 phosphate therefore has intrinsic 2-fold symmetry. Similarly, possible permutations in the $[-O-(PO_3)]^{2-}$ fragment required 3-fold overlap of the fragment over the template, and the resulting Na^+ distribution had a 3-fold symmetry.

The sodium distributions were transformed into pseudo electron densities using the previously described procedure.[36,37] The method involves a Fourier transformation of the 3-D distribution of Na^+ positions into their pseudo electron densities.[38] The densities were displayed and peaks fitted using the program O.[35] Sites of high densities represent sites of preferential sodium binding.

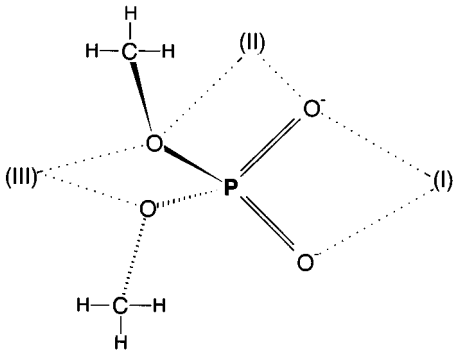

Figure 2. Dimethyl phosphate anion (DMP^-) and Na^+ starting positions. The Na^+ *ab initio* optimization started from positions I, II, and III, which are 2.33 Å from the neighboring oxygens and lie in the O═P═O (I), O═P−O (II), and O−P−O (III) planes, respectively.

2. *Ab Initio* Quantum Calculation. 2.1. Systems Studied. The use of *ab initio* techniques is still limited to small- and medium-sized molecules. Therefore, we were able to consider only carefully selected model systems. The smallest model systems were (di)hydrogen phosphate anions $H_2PO_4^-/HPO_4^{2-}$ and their sodium salts. These systems cannot represent the phosphodiester linkage in the DNA backbone, but they allowed us to consider larger basis sets, perform higher-level calculations, and compare these calculations to crystal structures. The phosphate geometry optimization started with parameters derived from crystal structures containing $H_2PO_4^-$. The sodium cation in NaH_2PO_4 was placed 2.33 Å from both partially charged oxygen atoms into the O═P═O plane.

Dimethyl phosphate anion $(CH_3-O-P(O_2)-O-CH_3)^-$ (DMP) and its sodium salt DMPNa (Figure 2) represent the smallest models containing the C−O−P−O−C phosphodiester linkage of the nucleic acid backbone which can still be studied by post-Hartree−Fock *ab initio* techniques. All optimizations of these systems started with phosphate geometry parameters derived from the crystal structures with phosphates of type $C_{sp^3}-O-P(O_2)-O-C_{sp^3}$ and charge −1 (labeled as standard in Table 1). The same starting geometry was used for the anion and its sodium salt. Three initial positions of the sodium cation were considered (Figure 2): (I) the symmetric position between both partially charged oxygens; (II) the position between the charged and substituted oxygens; (III) the symmetric position between both substituted ester oxygens. The Na^+ cation was in all cases located 2.33 Å from both neighboring oxygen atoms.

A larger system was needed to simulate more realistic interactions of the polyionic DNA backbone with metal cations. A fragment, $[CH_3-O-P(O_2)-O-(CH_2)_3-O-P(O_2)-O-CH_3]^{2-}$ (diPh; Figure 3), was built in a conformation imitating the canonical B-DNA as it has been established from fiber diffraction.[39] The fragment follows the main chain of the DNA backbone but does not include the branched deoxyribose atoms O4′, C1′, and C2′, or the atoms of a nitrogenous base. The $diPh^{2-}$ anion was complexed with either one Mg^{2+} cation or two Na^+ cations.

The isolated $diPh^{2-}$ fragment was fully optimized at the HF level, and the diPhMg and $diPhNa_2$ systems were optimized by using two different strategies: (1) All coordinates were fully optimized. (2) Only intermolecular coordinates of cations were optimized, and the intramolecular degrees of freedom of $diPh^{2-}$ were held rigid. These strategies were chosen not only for computational economy but also to maintain biologically relevant conformations in the studied fragments. The starting positions of Na^+ were the same as those used for calculation of DMPNa, and the starting positions of Mg^{2+} were analogous, but the distances from the nearest oxygen atoms were 2.03 Å.

(30) Hamilton, W. *Statistics in Physical Science;* Ronald Press: New York, 1964; pp 92−93.
(31) Clowney, L.; Jain, S. C.; Srinivasan, A. K.; Westbrook, J.; Olson, W. K.; Berman H. M. *J. Am. Chem. Soc.* **1996**, *118*, 509−518.
(32) Klyne, W; Prelog, V. *Experientia* **1960**, *16*, 521−523.
(33) *STATISTICA, A Program for Statistical Calculations*; StatSoft: Tulsa, OK, 1994.
(34) Cohen, D.; Vadaparty, K.; Dickinson, B. Efficient Algorithms for Distance Queries in Macromolecular Structure Databases. Technical Report No. ISL 95-25; University of Pittsburgh: Pittsburgh, 1995.
(35) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr.* **1992**, *A47*, 110−119.
(36) Schneider, B.; Cohen, D. M.; Schleifer, L.; Srinivasan, A. R.; Olson, W. K.; Berman, H. M. *Biophys. J.* **1993**, *65*, 2291−2303.
(37) Schneider, B.; Berman H. M. *Biophys. J.* **1995**, *69*, 2661−2669.
(38) Brünger, A. T. *X-Plor, Version 3.1. A System for X-Ray Crystallography and NMR*; Yale University Press: New Haven, CT, 1992.
(39) Chandrasekaran, R.; Arnott, S. *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids*; Springer-Verlag: Berlin, Germany, 1989.

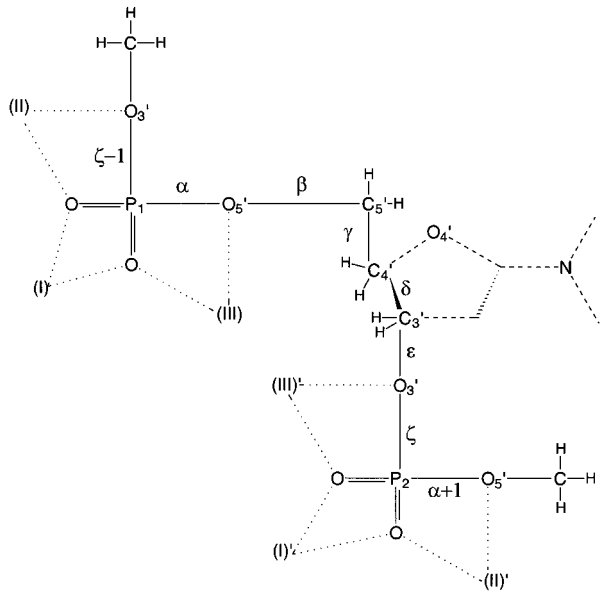

Figure 3. A model diphosphate compound (diPh) and starting positions for *ab initio* optimization of Na^+ and Mg^{2+} positions. The diPh was built from atoms connected by full lines in a conformation representing the fiber B-DNA backbone.[39] Two Na^+ cations were optimized starting from all possible combinations of positions I/I′, II/II′, and III/III′, which are analogous to the positions used in the DMPNa system. One Mg^{2+} was optimized starting from positions I, II, and III. Only some positions are stable after the optimizations; see Table 6.

2.2. Methods and Basis Sets. The systems studied were investigated at the Hartree-Fock *ab initio* level and also at a correlated level; the correlation energy was evaluated employing the second-order Møller−Plesset (MP2) perturbation theory. Geometries of the studied system were determined by gradient optimization. In most cases, all degrees of freedom were optimized using the first or second derivatives of the total energy, but only selected internal coordinates were optimized in the cases of the largest systems. All calculations were performed with the GAUSSIAN-92[40a] and GAUSSIAN-94[40b] sets of programs.

Various basis sets starting from split-valence (3-21G, 6-31G) through valence double-ζ to basis sets containing polarization functions (6-31G*, 6-31G**, D95**, D95(2d,p)) were employed. The $H_2PO_4^-$, NaH_2PO_4, DMP^-, and DMPNa systems were fully optimized at the HF and MP2 levels using the 6-31G* basis set. The diPh systems were optimized at the HF level by gradient optimization using the 6-31G basis set with polarization d-functions ($\alpha = 0.55$) on the phosphorus atoms.

Results and Discussion

The search for crystal structures containing the phosphate group resulted in 178 structures. A bibliography of these structures along with extended versions of Tables 1, 3, and 4 is deposited as Supporting Information. The average bond length and bond angle parameters observed in crystals are given in Table 1. Distributions of the torsion angles at the phosphodiester linkage are plotted in Figure 4. Contacts between the phosphate group and metal cations observed in crystals are summarized in Table 2, and the main interaction sites of Na^+ with the phosphate group are depicted in Figure 5. The results obtained by *ab initio* calculations on the dihydrogen and hydrogen phosphate systems are summarized in Table 3, on the dimethyl phosphate (DMP) systems in Table 4, and on the diphosphate model system diPh from Figure 3 in Table 5.

1. Crystal Data Analyses. 1.1. Phosphate Bond Geometry (Table 1). The phosphate group is in all cases arranged as a deformed tetrahedron. The average phosphate group with charges 0 and −2 has an approximate 3-fold symmetry along the P═O (for uncharged) or P−O (for −2) bonds. The phosphate with −1 charge has a 2-fold axis bisecting the two P═O bonds.

As is clear from Figure 1, the charge primarily determines phosphate stereochemistry since the uncharged group bears three substituents, the negatively charged group two substituents, and the twice charged group one substituent (Figure 1b,c,d). It is not, therefore, surprising that charge significantly influences the values of all bond parameters. The charge has the largest influence on the P═O and P−O distances which are elongated by 0.02−0.03 Å per each unit of negative charge and on the O═P═O and O═P−O angles. The first angle decreased by 4−6° on going from −1 to −2 phosphates (Table 1).

Within each charge group, the phosphate geometry is also influenced by two other factors: the number of carbon substituents and their chemical nature. The number of carbon substituents influences the P═O bonds and the O═P═O angles. Increasing the number of carbon substituents shortened the bond length by 0.02 Å and increased the angle by 2−3° per substituent.

The remaining bond parameters, i.e. P−O, C−O, O−P−O, and O═P−O, depend on charge and the chemical nature of the substituents. Four types of substitution proved to significantly alter the average values of these bond parameters: (1) acyclic

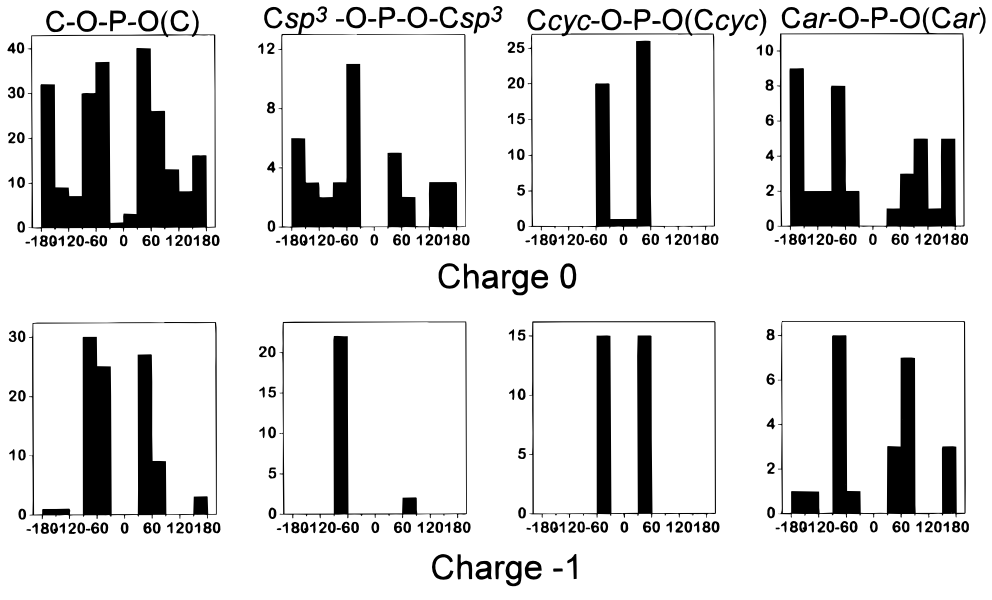

Figure 4. Histograms of the torsion angles C−O−P−O(−C) in uncharged and −1 phosphates.

Table 1. Bond Lengths (Å) and Valence Angles (deg) in the Phosphate Group As Observed in Crystal Structures[a]

bond and angle	phosphate charge	no. of carbons[b]	mean[c]	ESD[c]	*N*[c]	equal to standard[d]	*P*[d]
(C_{sp^3}—)O—P	**0 std**	**3**	**1.563**	**0.013**	**12**		
(C_{sp^3}—)O—P	**−1 std**	**2**	**1.595**	**0.015**	**16**		
(C_{sp^3}—)O—P	−1	1	1.598	0.009	22	Y	0.294
(C_{cyc}—)O—P	−1	2	1.599	0.013	44	Y	0.354
(C_{ar}—)O—P	−1	2	1.607	0.014	27	N	0.014
(H—)O—P	−1	0	1.565	0.009	64	N	0.000
(C_{sp^3}—)O—P	**−2 std**	**1**	**1.621**	**0.009**	**36**		
P═O	**0 std**	**3**	**1.449**	**0.011**	**34**		
P═O	**−1 std**	**2**	**1.485**	**0.013**	**16**		
P═O	−1	1	1.494	0.010	50	N	0.024
P═O	−1	0	1.503	0.008	64	N	0.000
P═O	**−2 std**	**1**	**1.514**	**0.008**	**108**		
C_{sp^3}—O	**0 std**	**3**	**1.447**	**0.019**	**12**		
C_{sp^3}—O	**−1 std**	**2**	**1.439**	**0.011**	**16**		
C_{cyc}—O	−1	2	1.441	0.011	44	Y	0.538
C_{ar}—O	−1	2	1.394	0.009	27	N	0.000
C_{sp^3}—O	**−2 std**	**1**	**1.433**	**0.012**	**36**		
(C_{sp^3}—)O—P—O(—C_{sp^3})	**0 std**	**3**	**103.3**	**2.6**	**12**		
(C_{sp^3}—)O—P—O(—C_{sp^3})	**−1 std**	**2**	**105.0**	**0.9**	**8**		
(C_{ar}—)O—P—O(—C_{ar})	−1	2	102.6	2.3	13	N	0.003
(C_{cyc}—)O—P—O(—C_{cyc})	−1	2	103.1	0.9	22	N	0.000
(C_{sp^3}—)O—P—O(—H)	−1	1	104.4	2.2	32	Y	0.346
(H—)O—P—O(—H)	−1	0	105.4	1.8	32	Y	0.382
(C_{sp^3}—)O—P═O	**0 std**	**3**	**115.2**	**1.7**	**12**		
(C_{sp^3}—)O—P═O	**−1 std**	**2**	**111.1/105.1**	**1.2/1.7**	**16**		
(C_{cyc}—)O—P═O	−1	2	109.4/107.6	1.2/1.0	44	N/N	0.000/0.000
(C_{ar}—)O—P═O	−1	2	110.8/105.8	1.2/2.1	27	Y/Y	0.434/0.240
(H—)O—P═O	−1	1	111.6/108.3	1.3/1.4	25	Y/N	0.216/0.000
(C_{sp^3}—)O—P═O	**−2 std**	**1**	**107.5/105.2**	**0.9/2.1**	**72/36**		
O═P═O	**−1 std**	**2**	**119.3**	**1.2**	**8**		
O═P═O	−1	1	116.9	1.4	25	N	0.000
O═P═O	−1	0	115.0	1.3	31	N	0.000
O═P═O	**−2 std**	**1**	**113.0**	**1.2**	**108**		
C_{sp^3}—O—P	**0 std**	**3**	**121.7**	**2.5**	**12**		
C_{sp^3}—O—P	**−1 std**	**2**	**120.1**	**1.1**	**16**		
C_{cyc}—O—P	−1	2	116.1	2.7	44	N	0.000
C_{ar}—O—P	−1	2	124.0	2.1	27	N	0.000
C_{sp^3}—O—P	**−2 std**	**1**	**119.1**	**2.0**	**36**		

[a] The standards defined in the Methods are in boldface type. The geometries of many more classes of phosphates are available as supporting information. [b] Number of carbon substituents attached to the ester oxygen(s). [c] The arithmetic mean of a parameter, estimated standard deviation (esd), and number of values in the sample (*N*). [d] A parameter mean is called different from the mean of the standard class when the probability of the Fisher−Behrens *t*-test is less than 0.05. The means were compared to the means of the standard classes separately for each charge.

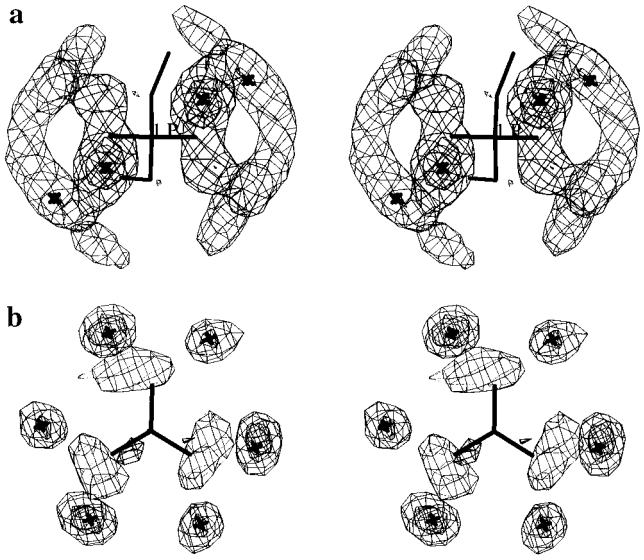

Figure 5. Stereo plots of Na^+ densities around the phosphate group with charges −1 (a, top) and −2 (b, bottom) derived from crystal data. The preferred Na^+ binding sites are indicated by crosses.

carbon in the sp^3 hybridization (labeled C_{sp^3}), (2) carbon in the sp^3 hybridization that is part of a four- or five-membered ring which rejoins the phosphate at the other ester oxygen (C_{cyc}, Figure 1f), (3) carbon in an aromatic ring (C_{ar}), (4) acyclic carbon in the sp^2 hybridization (C_{sp^2}), (5) hydrogen (H). Further subdivisions of these classes of structures did not significantly change the bond parameter values.

The most complex behavior was observed in the O═P—O angle. In phosphates with charge 0, a single mean value depends only on substituent type. In contrast, four O═P—O angles in −1 phosphates and three in −2 phosphates form two clusters with "large" and "small" values (Figure 6). The difference is most apparent for −1 phosphates where their respective mean values are 5° apart. In −2 phosphates, larger angles form 2/3 of the total population and the means differ by 2.5°. The difference becomes less pronounced for phosphates substituted by hydrogens because of their small volume and for cyclic phosphates probably because of steric restraints. For the larger means, charge reduces the angle by about 4° per each charge unit.

(40) (a) Frisch, J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C; Martin R. L.; Fox, D. J.; Defrees, J.; Baker J.; Stewart, J. P.; Pople J. A. *Gaussian 92*, Revision G.4; Gaussian Inc.: Pittsburgh, 1992. (b) Frisch, J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, M. A.; Ortiz, J. V.; Foresman, J. B.; Peng, C.Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin R. L.; Fox, D. J.; Binkley, J. S.; Defrees, J.; Baker J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C; Pople J. A. *Gaussian 94*, Revision B.3; Gaussian Inc.: Pittsburgh, 1994.

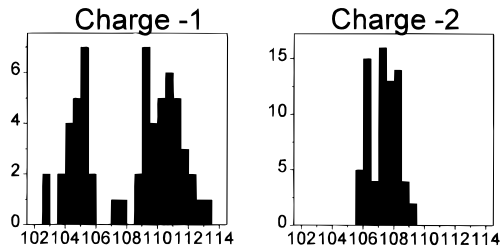

Figure 6. Histogram of O═P─O(─C_{ar}) bond angles in −1 and −2 phosphates with aromatic substitution.

The split of O═P─O angles in −1 and −2 phosphates is caused by steric repulsion between the partially charged oxygens and substituents of the other two oxygens in conformations which lack the C/H─O─P─O─C/H symmetry plane. It was concluded from 6-31G* MP2 model calculations on $H_2PO_4^-$ which showed that all four O═P─O angles remained the same only when the torsions were constrained to 180° so that the $(H_2PO_4)^-$ anion possessed the H─O─P─O─H symmetry plane.

The effect of substitution is best visible at the C─O and P─O─C parameters, but also at the P─O and O─P─O parameters (Table 1). Within each charge group, all these parameters acquire distinct values for different substituents. For instance, by replacing a C_{sp^3} with a C_{ar} atom, C─O shortens by 0.05 Å and P─O─C increases by 2−3°. A decrease of this angle by 4° after the $C_{sp^3} \rightarrow C_{cyc}$ substitution was most probably caused by the ring closure. The difference between the carbon and hydrogen substitution is largest for the P─O bond which is shorter for hydrogen phosphates by 0.02−0.04 Å but is also present for the O─P─O and O═P─O angles.

Our overview of the phosphate geometry extends earlier studies. An early work by Corbridge[41] does not discuss the geometries as a function of charge and substitution. An exhaustive and extremely useful survey of bond lengths by Allen *et al.*[42] lists several mean values for the phosphate bonds. The comparable means are very similar in both studies and deviate by about 0.005 Å, but due to an increase in the number of available crystal structures, this work presents more classes of phosphate bond parameters. The C_{sp^3} class, from which we excluded the nucleotide structures, agrees within 0.01 Å with the mean bond parameters derived for the DNA backbone.[1] The esds in Allen *et al.*[42] and in this work are almost identical. They are also similar to the esds of most bond parameters in fairly flexible ribose sugars[1] but are larger than the esds of parameters in very rigid nucleic acid bases.[31]

The sum of the six Y─X─Y angles in a regular tetrahedral group, XY_4, is 656.8°. The oxygen tetrahedrons from crystal phosphates deviate little from this value, generally by 0−3°. The largest deformation to 648° was observed for uncharged phosphates substituted by three aromatic carbons. This deformation cannot be explained by repulsion of bulky aromatic groups because the O─P─O angle, which would have to reflect such a repulsion, is actually smallest in aromatic phosphates and largest in hydrogen phosphates. In fact, the O─P─O angle has in all classes values smaller than 109.5°. On the other side, the O═P═O angles deviate to larger than tetrahedral values due to repulsion between the partially charged oxygens.

Baur[43] has described a phosphate group as a tetrahedron of oxygens in which the P atom is displaced from the centroid and much of the P···O bond length variation can be explained by off-center displacements of P. Over 80% of the P atoms are in the site with symmetry 1, and only a few phosphates display higher symmetries such as 222 or 23. In other words, phosphate geometries are locally deformed. The distortions of phosphate tetrahedrons have been correlated with hydrogen bond lengths[44,45] for crystals containing H_nPO_4. The distance between a phosphate oxygen and its hydrogen bond donor weakly correlates with P─O and O─P─O parameters. Souhassou *et al.*[46] have estimated the hydrogen bonding effects on the valence geometry of H_3PO_4 by comparing the crystal and *in vacuo* geometries. Parameters in the crystal, which are influenced by hydrogen bonding, are 0.04 Å longer for P═O and 0.02 Å shorter for P─O; the angles differ by 2−4°. The same effect was observed here when we compared geometries of $H_2PO_4^-$ and NaH_2PO_4 (Table 3) and is not therefore specific for hydrogen bonding interactions.

We did not attempt to relate the intramolecular bond distances and angles to hydrogen bonding patterns. After the classification of phosphate structures by charge and substitution, the remaining variation in the bond lengths (see esds in Table 1) is too little to discover principally important correlations with hydrogen bond distances. The bonding parameters depend primarily on the chemical environment and not on the nonbonding environment.

Another cumulative measure of the phosphate geometry─the sum of all P···O bond lengths─has been proposed by Cruickshank.[47a] The value of 6.184 Å derived for a limited set of −1 phosphates is surprisingly close to the values derived here. For instance, the C_{sp^3} class has the sum of the four O─P distances close to 6.160 Å for all three charge groups, and the largest deviation is observed for H_3PO_4 (6.103 Å). A constant sum of the four O─P distances has been observed also by Blessing.[48]

Cruickshank[47] and others[48,49] have discussed the role of d orbitals for phosphate structure and chemistry. The d orbitals are unoccupied but energetically accessible for mixing into the molecular orbitals. The phosphate geometry and bond strengths may be rationalized by a simple interplay between the interaction of the tetrahedral sp^3 orbitals of phosphorus with the p orbitals of oxygens and a partial donation of electrons from the oxygen free pairs to the phosphorus d_{z^2} and $d_{x^2-y^2}$ orbitals,[47a,48] but the d orbitals may just transfer electrons from near the bound atoms into the bond region.[47b]

1.2. Phosphate Torsion Angles. Distributions of the phosphate C─O─P─O(C) torsion angles are in Figure 4 plotted for two groups of phosphate structures, with charge 0 and three carbon substituents, and with charge −1 and two carbon substituents. In order to separate the steric effects of various substituents, only those torsions were studied in which both carbons of C─O─P─O─C fragments were of the same type, acyclic C_{sp^3}, cyclic C_{sp^3}, or aromatic.

In all phosphates with charge −1, the most populated are both *sc* regions. The dominant *−sc*/*−sc* conformation in the C_{sp^3} class can be explained by the fact that most of these compounds are dinucleotides forming a right-handed minihelix. The mean value of the tight *−sc* cluster of −71° (esd 6°) is close to the means of the phosphodiester linkage torsions α and ζ in A-DNA oligonucleotides.[2]

(41) Corbridge, D. E. C. *The Structural Chemistry of Phospates*; Elsevier: New York, 1974.
(42) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. *J. Chem. Soc., Perkin Trans. 2* **1987**, S1−S19.
(43) Baur, W. H. *Acta Crystallogr.* **1974**, *B30*, 1195−1215.
(44) Ichikawa, M. *Acta Crystallogr.* **1987**, *B43*, 23−28.
(45) Ferraris, G.; Ivaldi, G. *Acta Crystallgor.* **1984**, *B40*, 1−6.
(46) Souhassou, M.; Espinosa, E; Lecomte, C.; Blessing, R. H. *Acta Crystallogr.* **1995**, *B51*, 661−668.
(47) (a) Cruickshank, D. W. *J. Chem. Soc.* **1961**, 5485−5504. (b) Cruickshank, D. W. *J. Mol. Struct.* **1985**, *130*, 177−191.
(48) Blessing, R. H. *Acta Crystallogr.* **1988**, *B44*, 334−340.
(49) Moss, G. R.; Souhassou, M.; Blessing, R. H. Espinosa, E; Lecomte, C.; *Acta Crystallogr.* **1995**, *B51*, 650−660.

Cyclic phosphates are sterically restricted to the +*sc*/−*sc* (or −*sc*/+*sc*) conformation due to the ring closure requirement. Other torsion regions are precluded because the five- or six-member P−O−$(C)_2$−3-O ring cannot adopt an *ap* arrangement and the *sp* region is energetically inaccessible. The C−O−P−O torsion angles, labeled α and ζ in nucleotide chemistry, are negatively correlated such that ζ = −α. The mean of the −*sc* conformation of −51° (esd 3°) is significantly lower than in the C_{sp^3} class. The +*sc* conformation has a mean of 50°. The uncharged cyclic phosphates showed the same behavior with the means of −43° and 43° (esd 9° and 8°) and the same negative correlation between both torsions. The lower torsion means of neutral cyclic phosphates are caused by additional steric restrictions imposed on the ring by the third substituent.

The C_{ar}−O−P−O(C_{ar}) torsions have more complicated distributions. In negatively charged phosphates, they prefer the ±*sc* regions with similar mean values (−71° and 72°) as the −1 C_{sp^3} class, but the clusters are much looser (esds 10° and 14°). The aromatic substituents were also found in extended conformations with torsions in the *ap* and *ac* regions (mean −176°, esd 30°). The existence of extended conformations may be caused by either steric repulsion of bulky aromatic groups in ±*sc*/±*sc* conformations or better packing of structures in which these bulky groups are in a more linear arrangement. The −*sc*/−*sc* and +*sc*/+*sc* clusters are still most populated, and −/+ or +/− combinations were not observed. An *ap* torsion always combines with ±*sc* but never with another *ap*.

In uncharged phosphates with aromatic substituents, the planar *ap* cluster (mean −169°, esd 20°) is more populated than both *sc* clusters (means −68° and 82°, esds 18° and 20°). The tendency of aromatic substituents to adopt more extended conformations is therefore more pronounced in uncharged than in −1 phosphates. Uncharged phosphates with three carbon substituents have six C−O−P−O(−C) torsion angles. Given the tetrahedral arrangement of the phosphate group, the maximum number of energetically preferred *sc* torsion angles is four, and two torsions have to adopt less stable *ap* torsions. The observed ratio between *ap* and *sc* conformations in aromatic phosphates is higher than 2:6, but exact evaluation of an excess of *ap* conformers is difficult because the third substituents can be C_{ar}, C_{sp^3}, or C_{sp^2}, or even H.

The C_{sp^3}−O−P−O(−C_{sp^3}) torsion distribution in uncharged phosphates is similar to that of the uncharged aromatic class with means of −67° and 76° for ±*sc* (esds 25° and 23°) and −173° (esd 26°) for *ap*. The *ap* cluster represents approximately expected 2/6 of the total population. Very large esds indicate broad, poorly clustered distributions. The scattergram of two torsions from the C_{sp^3}−O−P−O−C_{sp^3} moiety showed dispersed clusters in the −*sc*/−*sc*, ±*sc*/*ap*, and even *ap*/*ap* conformations.

The only carbon substituent of −2 phosphates forms three C−O−P═O torsion angles (distributions not shown). Because these angles are not independent, we calculated the mean value for one torsion within 0° and 120°. Not surprisingly, most of these torsions are close to +60° with a mean of 55° (esd 13°).

All surveys of experimentally and theoretically determined phosphate conformations have been based on model systems classified here as the C_{sp^3} class of −1 phosphates, and most work has been done on DMP^-. The −*sc*/−*sc* arrangement has been identified as the most populated in nucleotides since the early work on their crystal structures.[50−52] Other experimental techniques have also predicted *sc* regions as the most populated.

Table 2. Interaction of Metal Cations with the PO_4 Group in Crystals

	total number of analyzed				contacts from cation to[b]		
system	structures	cations	H_2O	distant M^+ [a]	PO_4	H_2O	other
$(PO_4)^-/Na^+$	8	17	21	3	29	24	30
$(PO_4)^{2-}/Na^+$	15	41	85	24	18	134	54
$(PO_4)^-/K^+$	5	8	1	3	17	1	17
$(PO_4)^{2-}/K^+$	2	4	10	2	2	10	12
$(PO_4)^-/Mg^{2+}$	3	6	6	0	20	18	14
$(PO_4)^{2-}/Mg^{2+}$	6	7	53	0	12	31	13

[a] Number of cations not coordinated by a phosphate defined as having a distance greater than 3.00 Å from any of the phosphate atoms. [b] Number of contacts with distances less than 3.00 Å between a cation and a specified group.

Despite a significant flexibility of the C−O−P−O−C moiety in solution, ^2H and especially ^{31}P NMR experiments have localized −*sc*/−*sc* (and +*sc*/+*sc*) as the most stable conformations[53−55] although *sc*/*ap* and possibly *ap*/*ap* have also been detected.[53] The same conformers have been identified as the most stable by IR and Raman spectroscopy,[4−6] reviewed in a paper by Thomas and Tsuboi.[56] Extended conformational states (±*sc*/*ap* and *ap*/*ap*) of DMP^- have been observed by depolarized Raleigh scattering.[57] *Ab initio* calculations of DMP^- by Newton[3] and other authors[9,10,18,20] recognize −*sc*/−*sc* conformers as the most stable but also find ±*sc*/*ap* conformations energetically accessible. Monte Carlo simulations of DMP^- in aqueous solution[58−60] have found that the stabilities of DMP^- conformations decrease in the following order: −*sc*/−*sc*, ±*sc*/*ap*, *ap*/*ap*. Approximate 2-fold symmetry of the most stable −*sc*/−*sc*, (+*sc*/+*sc*) conformation of DMP is also reflected in its experimental charge distribution.[61] It is similar at both ester and at both charged oxygens, but the two kinds of oxygens are significantly different.

1.3. Interactions with Metal Cations. The occurrence of Na^+, K^+, and Mg^{2+} in well-refined phosphate crystal structures was sufficient to study interactions between these metals and phosphate. The statistics presented for these cations in Table 2 showed that most of the structures contained Na^+ counterions. The numbers of water molecules interacting directly with phosphate oxygens show that −1 phosphates contain one water or less per cation and are much less hydrated than −2 phosphates, where this ratio is about two in Na^+ and K^+ and even higher in Mg^{2+} salts.

The numbers of cations not directly coordinated by the phosphate oxygens suggest a more surprising distinction between interactions of Na^+ with −1 and −2 phosphates (Table 2, fifth column). Of the total 17 Na^+ ions in −1 phosphates, 14 are coordinated directly by phosphate oxygens so that cations and anions compensate for their charges by a mutual close contact. In twice charged phosphates, only 17 of 41 Na^+ ions are coordinated by phosphate oxygens even when they should attract their counterions more forcefully. Most of the −2 phosphate/

(50) Sundaralingham, M. *Biopolymers* **1969**, *7*, 821−860.
(51) Yathindra, N.; Sundaralingham, M. *Biopolymers* **1973**, *12*, 297−314.
(52) Kim, S. H.; Berman, H. M.; Seeman, N. C.; Newton, M. D. *Acta Crystallogr*. **1973**, *B29*, 703−710.
(53) Gorenstein, D. G. In *Phosphorus-31 NMR,Principles and Applications*; Gorenstein, D. G., Ed.; Academic: New York, 1984; Chapter 1.
(54) Nikonowicz, E. P.; Meadows, R. P.; Fagan, P.; Gorenstein, D. G. *Biochemistry*, **1991**, *30*, 1323−1334.
(55) Akutsu, H.; Nagamori, T. *Biochemistry* **1991**, *30*, 4510−4516.
(56) Thomas, G. J., Jr.; Tsuboi, M. *Adv. Biophys. Chem.* **1993**, *3*, 1−69.
(57) Garrigou-Lagrange, C.; Bouloussa, O; Clement, C. *Can. J. Spectrosc.* **1976**, *21*, 75.
(58) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc*. **1985**, *107*, 2229−2239.
(59) Jayaram, B.; Mezei, M.; Beveridge, D. L. *J. Comput. Chem*. **1987**, *8*, 917−942.
(60) Jayaram, B.; Mezei, M.; Beveridge, D. L. *J. Am. Chem. Soc*. **1988**, *110*, 1691−1694.
(61) Klooster, W. T.; Craven, B. M. *Biopolymers* **1992**, *32*, 1141−1154.

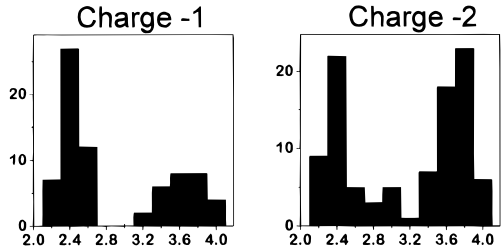

Figure 7. Histogram of distances between charged oxygens in −1 and −2 phosphates.

Na^+ contacts are mediated by waters as there are 134 contacts between Na^+ and water and only 18 contacts between Na^+ and a phosphate.[62] Contacts of K^+ and Mg^{2+} indicate the same trend.

The difference between solvation of −1 and −2 phosphates is illustrated by the distribution histograms in Figure 7. While −1 phosphates have most Na^+ ions in the first solvation sphere and the gap between the first and second shells of cations is wide, −2 phosphates have more populated the second shell with many other Na^+ ions even further and the first and second shells partially overlap. A typical solvation of the phosphate group by Na^+ can therefore be summarized as $Ph^--Na^+(W)-X$ for −1 phosphate and $Ph^{2-}-W-Na^+(W)-W-X$ for −2 phosphate where W is a water molecule, $Na^+(W)$ a hydrated sodium cation, and X a polar group. An analysis of distributions of other cations around −1 and −2 phosphates will allow one to decide whether it is a unique property of Na^+ or a general behavior of metal cations.

Numbers of Na^+ ions solvating −1 and −2 phosphates (Table 2) allowed the Na^+ spatial distributions to be determined. Both Na^+ distributions plotted at the same density levels of 4- and 8σ demonstrate that Na^+ positions around the phosphate group are not random and that they are located in well-defined sites. The −1 distribution (Figure 5a) was drawn with the 2-fold symmetry axis approximately perpendicular to the paper plane so that the left and right sides are equivalent. Each charged oxygen is solvated by two dominant Na^+ sites which are indicated by crosses. The major site has a significantly higher density than the minor site (12σ versus 8σ). Both sites are 2.40 Å from the charged oxygen, their $Na^+\cdots O=P$ angles are about 122°, and the $Na^+\cdots O=P-O(-C)$ torsion angles are nonplanar, 30° for the higher site and −118° for the lower site. A notable feature of the distribution is that there is no Na^+ population in the O−P−O plane and along the O=P vectors. A low bagel-shaped density indicates that metals can bind in the whole circular region around the charged oxygens. This observation is in agreement with a charge density study by Moss *et al.*[49] in which the authors have stated that "The (P=)O lone-pair density is almost circularly symmetric about the P=O axis..."

The Na^+ distribution around −2 phosphates was depicted with the 3-fold symmetry axis perpendicular to the paper plane and charged oxygens facing above the page (Figure 5b). The substituted ester oxygen is therefore below the paper hidden behind the phosphorus atom. As in the −1 distribution, each charged oxygen is solvated by two sodium sites. The major and minor sites have the same ratio of their densities (12- and 8σ), and also their $Na^+\cdots O$ distances (2.22/2.45 Å) and $Na\cdots O=P$ angles (136°/124°) are similar as in −1 phosphates. A shortening of the $Na^+\cdots O$ distances in −2 phosphates is not significant. The $Na^+\cdots O=P-O(-C)$ torsions are 70° for the major site and −65° for the minor site.

(62) Note a difference between the number of analyzed cations and the number of their contacts: One Na^+ can make two contacts to phosphate oxygens, one contact to a water, and two contacts to other nucleophilic groups such as a carboxyl group.

The distances between the Na^+ sites and the charged oxygens are only slightly longer than the Na^+−water internuclear distance reviewed by Marcus.[63] The mean distances between K^+, Mg^{2+}, Ca^{2+} and charged phosphate oxygens (2.76, 2.06, and 2.42 Å; esds about 0.10 Å) determined from statistically less representative samples of crystal structures (Table 2) are also in excellent agreement with data presented by Marcus.[63] This supports our observation that distances of a cation to its closest neighbors do not depend on the neighbor's charge, and this further indicates that differences between equilibrium distances in crystal and solution environments are not large.

Alexander *et al.*[27] have determined preferred metal sites around the phosphinyl group ($-PO_2^--$) from the crystal data. The preferred metal sites around the $(-PO_2-)^-$ and $(-O-PO_2-O-)^-$ groups qualitatively agree; in both cases there is no density in the O−P−O plane, and the main metal site is located in a slightly nonplanar position between the charged oxygens. Quantitative parameters are, however, hard to compare because the distribution around the phosphinyl group has been based on only 34% of phosphate structures and on a wide variety of metals from Na^+ to U^+ and metal distances from the phosphinyl oxygens were limited to 2.5 Å. The $Na^+\cdots O=P$ angle of 141° reported by Alexander *et al.*[27] differs from the values derived here (about 120°). The difference could arise because about half of their sample[27] is formed by transition metals which are known to bind more covalently than electrostatically binding alkali metals. Some bias of the metal densities has also been introduced by symmetrization of the data by two mirror planes,[27] which is not possible in the case of the phosphate group because it has only one 2-fold axis in its most populated *sc*/*sc* conformation.

In a 150 ps molecular dynamics simulation, Laughton *et al.*[21] observed an indication that Na^+ counterions are associated with just one of the two charged phosphate oxygens and that they are rarely found along the O=P=O bisector. This is in accord with the positions of the preferred Na^+ sites derived here.

2. Ab Initio Calculations. 2.1. Hydrogen Phosphates. Model calculations of hydrogen phosphates (Table 3) revealed that at least one set of polarization functions must be put on non-hydrogen atoms for the adequate description of the phosphate geometry. This decreased bond lengths by about 0.1 Å and improved their agreement with the experimental values; a second set of polarization functions decreased bond lengths by a much lesser degree. Valence angles are less sensitive to addition of polarization functions and change mostly by less than 1°. Change from the HF to MP2 level increased bond lengths by about 0.04 Å and did not change valence angles significantly.

A cation has an important effect on bond lengths and angles. For both $(H_2PO_4)^-$ and $(HPO_4)^{2-}$, the presence of Na^+ reduced the P−O bond length by about 0.04 Å per each compensated charge while the P=O bond was lengthened. For $(H_2PO_4)^-$, the O=P=O angle was reduced significantly (by about 10°), and the O=P−O and O−P−O angles slightly increased, while for $(HPO_4)^{2-}$, the largest change was observed in the O=P−O angle.

Optimization of sodium positions in NaH_2PO_4 started from the symmetric position between the charged oxygens. The comparison of MP2/6-31G** geometries of NaH_2PO_4 with the average crystal geometries of the dihydrogen phosphate anion (Table 3) shows that the theoretical P−O bond length is still too long, while the angles agree well. Further extension of a basis set is required to reach a better agreement with experimental data. Due to compensation of errors, the best agreement

(63) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475−1498.

Table 3. Bond Lengths (Å) and Valence Angles (deg) in (Di)hydrogen Phosphate Anions and Their Sodium Salts Determined by *ab Initio* Calculations at the MP2 Level[a]

		bond lengths			valence angles		
system	basis set	O—P	P═O	Na—O	O═P—O	O—P—O	O═P═O
HPO_4^{2-}	6-31G*	1.781	1.545/1.538		102.3/100.4		116.4/114.2
Na_2HPO_4	6-31G*	1.669	1.591/1.542	2.24	107.9/102.0		117.8/109.9
$H_2PO_4^-$	6-31G	1.783	1.609		109.6/104.8	99.5	125.5
$H_2PO_4^-$	6-31G*	1.680	1.505		108.0/105.7	100.8	125.9
$H_2PO_4^-$	6-31G**	1.678	1.505		108.2/105.6	100.6	125.7
NaH_2PO_4	6-31G*	1.639	1.522	2.27	110.9/107.5	103.6	115.7
NaH_2PO_4	6-31G**	1.638	1.522	2.27	110.8/107.6	103.7	115.7
$H_2PO_4^-$	crystal	1.565	1.503	2.40	111.3/106.6	105.4	115.0
$H_2PO_4^{2-}$	crystal	1.591	1.512	2.22	107.8/105.2		112.0

[a] The results are compared to the crystal values determined for the $[(H)—O—PO_2—O—(H)]^-$ and $[(H)—O—PO_2—O—(H)]^{2-}$ fragments (see also Table 1). A more complete version of the table was deposited as Supporting Information.

Table 4. Relative Stabilities of the Dimethyl Phosphate Anion (DMP^-) and Its Sodium Salt[a]

	Na···O	C—O—P—O(—C) Torsion (deg)		
system	distance (Å)	start	optimized	ΔE[b]
DMP^-		70/70	70/70	0
DMP^-		70/−70	90/−91	2.50
DMP^-		70/180	71/162	0.96
DMP^-		180/180	180/180	2.93
DMPNa(I)	2.27	70/70	69/69	0
DMPNa(I)	2.27	70/-70	85/−85	3.27
DMPNa(I)	2.26	70/180	64/166	1.34
DMPNa(I)	2.24	180/180	179/−179	3.59
DMPNa(II)[c]	2.16	70/70 or 180/70	72/73	7.24
	2.44			
DMPNa(II)[c]	2.16	70/−70 or 70/180	70/156	6.35
	2.48	or 180/180		
crystal	2.40		−71/−71	

[a] In both cases, the *sc*/*sc* conformation is the most stable. All degrees of freedom were optimized at the MP2 level using the 6-31G* basis set. Sodium positions I and II are defined in Figure 2. The results are compared to the crystal values determined for the $[(C_{sp^3})—O—PO_2—O—(C_{sp^3})]^-$ class of phosphates (see also Table 1). A more complete version of the table was deposited as Supporting Information. [b] Destabilization energy (kcal/mol) relative to the fully optimized DMP^- *sc*/*sc* for the charged systems and relative to the fully optimized DMPNa(I) *sc*/*sc* for the sodium salts. [c] A Na^+ position asymmetric relative to the phosphate-induced different values of some bond parameters. The first values are for the oxygen atoms close to Na^+ and the second ones for the oxygens far from Na^+.

with experimental data was obtained at the HF level with a medium-sized basis set such as 6-31G*; this level was therefore used for the largest model system diPh.

2.2. Dimethyl Phosphates (DMP). All conformations presented in Table 4 were fully optimized at the MP2 level. The geometric behavior of the DMP systems is, in many respects, similar to that of hydrogen phosphate systems: the adequate description required the use of polarization functions, the bond lengths increased and the angles decreased upon passing from the HF level to the MP2 level, and an addition of Na^+ shortened P—O, elongated P═O and C—O bonds, and significantly reduced the O═P═O angle.

The most stable conformation of DMP^- and of DMPNa calculated at the MP2/6-31G* level has the $+sc/+sc$ (or by symmetry $-sc/-sc$) conformation. This is in accord with calculations of other authors performed at the HF level.[3,9,10,18] The energetic differences between different conformations are relatively small,[18] within 3 kcal/mol. The *ap*/*ap* conformations of DMP^- and DMPNa(I) represent local minima, but they are the least stable. Even the $+sc/-sc$ conformation with a tight C···C nonbonding contact is slightly more stable than the *ap*/*ap* arrangement. It further emphasizes the importance of the anomeric (*gauche*) effect[64−66] for the stereochemistry of the phosphate group.

Some parameters depend on the O—P—O—C torsion angles. The O—P—O angle diminishes by 4° per change from an *sc* conformation to an *ap* conformation. The same dependence has been observed at the HF level by Gorenstein *et al*.[9,10] Bond parameters generally acquire their extreme values in the energetically least stable *ap*, *ap* conformation.

The O═P—O angles also depend on the C—O—P—O torsion angles. In agreement with crystal geometries (Table 1), the angles have two large and two small values in all conformations lacking the C—O—P—O—C plane of symmetry. On the other side, the calculated correlation between values of the O═P—O bond angles and the C—O—P—O torsions were not observed in the analyzed crystal data. This probably arose because effects imposed by substituents of the C—O—P—O—C moiety and by various crystal environments concealed the correlation.

Two stable configurations were found for the $DMP^-—Na^+$ system (Table 4). In position I, Na^+ is located symmetrically between two partially charged oxygens, while in position II, it is between one charged oxygen and one ester oxygen (Figure 2). Position III could not be optimized because Na^+ shifted to the site II. Na^+ prefers the symmetric position I over the asymmetric position II by at least 3 kcal/mol, and the global minimum is DMPNa(I) *sc*/*sc*. In both configurations, Na^+ virtually lies in the O═P═O (position I) or O═P—O (position II) plane, and the Na—O═P angles are close to 90°. Na^+ changed the DMP^- internal geometry similarly to the way it changed the $H_2PO_4^-$ and HPO_4^{2-} geometries. The largest effects were observed at the O—P bond and O═P═O angle.

Na^+ in the asymmetric configuration (II) induced significant changes in the phosphate geometry. Potentially equivalent distances and angles can become significantly different so that the two P—O bonds differ by 0.08 Å and the O═P—O angles by 10°. When Na^+ was moved from position I to position II in the stable *sc*/*sc* conformation, the O═P═O angle changed by 10° and the smaller of the O═P—O angles by 8°.

The conformations of DMP^- and DMPNa(I) are very similar, and their order of stability is the same for both systems, with *sc*/*sc* as the most stable and *ap*/*ap* as the least stable. Considering the symmetry of dimethyl phosphate in these systems, their conformational maps have stationary points in nine regions (four $\pm sc/\pm sc$, two $\pm sc/ap$, two $ap/\pm sc$, and one *ap*/*ap*). On the other hand, the symmetry of DMPNa(II) is reduced, and its conformational map has only six stationary points, with $sc/-sc$, $-sc/sc$ and *ap*/*ap* missing. The Na^+(II) position is stable only when attached to an ester oxygen which is a part of the O—P—O—C torsion in a $+sc$ conformation. When optimization

(64) Lemieux, R. U. *Pure Appl. Chem*. **1971**, *25*, 527−548.
(65) Wolfe, S. *Acc. Chem. Res*. **1972**, *5*, 102−111.
(66) Radom, L.; Hehre, W. J.; Pople, J. A. *J. Am. Chem. Soc*. **1972**, *94*, 2371−2381.

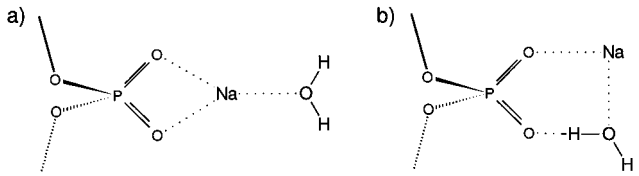

Figure 8. Structural formulas of hydrated Na^+ and Li^+ dihydrogen phosphates. For both cations, structure a was more stable by more than 5 kcal/mol (calculated at the HF/3-21G* level).

of DMPNa(II) started with this torsion angle in the *ap* region, it moved into the +*sc* region while the other torsion angle remained in either the +*sc* or *ap* region. The *sc* conformation is therefore stabilized relative to *ap* when Na^+ is in the asymmetric position II. In DMPNa(II), the relative stability of *sc/sc* and *sc/ap* is reversed, probably due to a short contact between the cation and a methyl hydrogen. The asymmetry of the cation position also shifts a +*sc/ap* conformation into the +*sc/ac* region. The *sc/*−*sc*, −*sc/sc*, and *ap*/*ap* conformations, which are unstable in DMPNa(II), are not observed in known DNA helical conformations.[2] In this context, a cation in an asymmetric position helps to stabilize the helical arrangement of the DNA backbone.

Comparing the crystal data for the $(C_{sp^3}{-}O{-}P{-}O{-}C_{sp^3})^-$ class of phosphate structures with the MP2 calculations of the DMPNa structure revealed similar discrepancies as in the case of H_2PO_4: the calculated bond lengths are systematically longer by about 0.03 Å, while the valence angles are close to the experimental values. The HF results are again closer to the crystal values due to compensation of errors.

Calculated intermolecular Na···O distances (Table 4) are shorter by 0.15 Å than the average experimental value. A disagreement between the most stable theoretical configuration DMPNa(I) and the experimental Na^+ distribution is, however, qualitative. While the theoretical site is located symmetrically between the charged oxygens in the O═P═O plane, the main experimental site lies out of this plane and is shifted closer to one of the charged oxygens. Moreover, the experimental distribution shows no sodium cations along the O═P═O bisector which is predicted to be the most favorable binding site for Na^+. The second ordered Na^+ site is even further from the Na(I) position. The Na(II) position does not have any close counterpart in the experimental distribution because both crystal ordered sites are attached only to the charged oxygens. Only two lobes of relatively low density sticking out of the bagel-shaped Na^+ locus (Figure 5a) can be compared to the Na(II) position.

Larger discrepancies between geometries calculated at a reliable MP2 level and their crystal counterparts may reflect flaws in a theoretical model rather than inadequacies in its description. In the case of charged systems such as a phosphate group, it may become necessary to include other parts of the environment than just a counterion to respect the key components of the condensed phase. Especially water molecules and other highly polar and polarizable particles can modify geometric and energetic characteristics of the central anion. Most crystal structures are hydrated with many water molecules in close proximity to both cation and phosphate groups.

Specifically, the discrepancy between the sodium sites suggested by theoretical calculation and by crystal structures can only be resolved by using theoretical models which contain water or other polar molecules. Therefore, we investigated two monohydrated models, $MH_2PO_4{-}H_2O$, where M^+ is either Li^+ or Na^+ (Figure 8). The full gradient optimization at the HF/3-21G* level showed that the cyclic form b is more stable for Li^+ by 5.8 kcal/mol and for Na^+ by 7.2 kcal/mol. In the cyclic form, the cation position qualitatively agrees with the experimentally determined Na^+ site even when there are still significant discrepancies: The $Na^+{\cdots}O$ distance is 2.08 Å, the $Na^+{\cdots}O{=}P$ angle 107°, and the $Na^+{\cdots}O{=}P{=}O$ torsion 0°. The question of whether models including more water molecules will bring theoretical models closer to the crystal distribution is being studied in our laboratory.

Table 5. Starting and Optimized Bond and Torsion Angles (deg) of the Diphosphate $diPh^{2-}$ and $diPhNa_2$ Model Systems[a]

angle definition	start	optimized $diPh^{2-}$	optimized $diPhNa_2$
O5′—P—O3′	101.4	99.3	100.6
C5′—O5′—P	119.1	121.0	125.3
P—O3′—C3′	119.0	123.1	126.4
O═P—O	109.6	107.7/106.1	112.9/110.8
O═P═O	116.1	122.4	110.1
$\zeta - 1$ (C3′—O3′—P1—O5′)	−157	−142	−177
α (O3′—P1—O5′—C5′)	−41	−80	−57
β (P1—O5′—C5′—C4′)	135	156	149
γ (O5′—C5′—C4′—C3′)	37	65	58
δ (C5′—C4′—C3′—O3′)	139	−178	174
ϵ (C4′—C3′—O3′—P2)	−133	−97	−92
ζ (C3′—O3′—P2—O5′)	−157	−72	−177
$\alpha + 1$ (O3′—P2—O5′—C5′)	−41	−65	−54

[a] The chemical formula and atom and torsion angle labels are in Figure 3. All degrees of freedom were fully optimized at the HF level with the 6-31G basis set and polarization d function ($\alpha = 0.55$) added to both phosphorus atoms. The starting bond distances and angles were taken from this study and Allen *et al.*[42] and torsion angles from the fiber B-DNA structure.[39]

An analogous preference for the asymmetric position of a cation has been observed for complexes of $H_2PO_4^-$ with hydrated Mg^{2+} and Ca^{2+} by Deerfield and Pederesen.[5] The origin of stabilization of the asymmetric position over the symmetric position between the charged oxygens is in the larger stabilization of Mg^{2+} by a water molecule than by the other charged oxygen.

Ab initio calculations of interactions between DMP^- and metal cations not considering water or other polar species are deemed to find the axial symmetrical arrangement as the most stable. Unfortunately, even models which do consider waters can fail to discover the most stable circular arrangement (Figure 8b). The $M^+{-}DMP^-{-}H_2O$ system has been assumed as the most stable and held in the C_{2v} linear arrangement[67] as in Figure 8a. A model of a hydrated nucleotide with Na^+ has also been used by Kim and LeBretton[68] to explain the mechanism of guanine methylation, but details about the studied structures supplied by the authors do not allow the determination of whether Na^+ was more stable in the symmetric or crystal-like position.

Alexander *et al.*[27] optimized the LiH_2PO_4 (HF/3-21G*) system and reported an asymmetric position as the most stable. This statement is, however, puzzling. We reproduced the reported minimum (the $Li{\cdots}O{=}P$ angle, 142°) which approximately agrees with the position derived from the crystal data, but the barrier for this minimum is only on the order of 0.1 kcal/mol, and the other, symmetric, minimum is deeper by 19 kcal/mol. This shallow asymmetric minimum was found neither using a larger basis set, 6-31G*, nor for Na^+.

2.3. Model Diphosphate Compound (diPh). The $diPh^{2-}$ system (Figure 3) and its sodium and magnesium salts were constructed to model DNA backbone−cation interactions. Full HF optimization of $diPh^{2-}$ proved that the bonding parameters

(67) Marynick, D. S.; Schaefer, H. F., III *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3794−3798.
(68) Kim, H. S.; LeBretton, P. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3725−3729.

Table 6. Geometry and Relative Energy of the $diPh^{2-}$ Complexed with Na^+ and Mg^{2+} Cations[a]

starting positions of Na^+	optimized positions	ΔE[b]
I + I′, II + I′, I + III′, or III + III′	I + I′	0
III + I′	III asym[c] + I′	7.1
II + II′	I + II′ asym[c]	8.4
III + II′	III asym + II′ asym[c]	15.1
starting position of Mg^{2+}	**optimized positions**	**ΔE[b]**
I or II	I	0
III	III asym[c]	11.5

[a] For a description of the starting positions of the cations, see Figure 3. Computed at the HF level with the 6-31G basis set and a polarization function at both P atoms. Only the positions of the cations were optimized. [b] Destabilization energy (kcal/mol) relative to the most stable partially optimized structure of $diPhNa_2$ or diPhMg. [c] Na^+ is closer to the partially charged oxygen atom.

have similar values as in DMP^-. The bond distances were within 0.01−0.02 Å from the HF values for DMP^-, and the bond angles deviate by 2−4°, with the largest deviations observed for the C−O−P angle.

The starting conformation of an idealized "fiber" B-DNA[39] did not change considerably after the optimization of either $diPh^{2-}$ or $diPhNa_2$ (Table 5). The average root mean square deviation between atom positions before and after optimization of $diPh^{2-}$ is 0.81 Å, and the torsion angles remained within limits observed in B-DNA[2]. Optimization of diPhMg induced a severe bend in the original B-DNA-like structure because Mg^{2+} formed a bridge between both negatively charged phosphate groups.

In order to investigate the backbone−cation interactions the conformation of B-DNA-like diPh was held rigid, and only cation positions were optimized. For two Na^+ cations, all combinations of the three positions I, II, and III were used as starting configurations (Figure 3). The optimized positions are summarized in Table 6. Surprisingly, these positions are virtually identical with Na^+ positions around DMP^-. This suggests that a model system containing just one phosphate group is sufficient to model the phosphodiester linkage of the DNA backbone. The most stable (I + I′) arrangement has both cations located symmetrically between the charged oxygens in the O═P═O plane (Na···O distance, 2.3 Å). The largest deviation (10°) from this plane was observed in the (I + II′) arrangement, while the largest deviation (18°) from the O═P—O plane came from the least stable (III + II′) arrangement. All Na···O═P angles are close to 90°.

The optimization of Mg^{2+} in positions I, II, and III converged in the symmetric and planar configuration I (Mg···O, 2.0 Å; Mg···O═P, 94°) or the asymmetric and energetically much higher configuration (III) in which Mg^{2+} is closer to the charged oxygen.

As in the case of DMPNa, none of the stable calculated configurations of $diPhNa_2$ or diPhMg can be identified with either of the experimental Na^+ sites. We therefore conclude that a much larger and, hence, a more "realistic" model did not bring theory and experiment closer to each other. In order to effectively simulate the behavior of cations in the condensed phase, a model must contain other polar particles in addition to the ionic pair.

Acknowledgment. This work was supported by grants from the Grant Agency of the Czech Republic to B.S. (204/95/0036) and for purchase of the CSD (203/96/0111). We thank the reviewers for stimulating comments and Julie Woda for language corrections.

Supporting Information Available: A bibliography of the 178 structures obtained in the search for crystal structures containing the phosphate group and extended versions of Tables 1, 3, and 4 (18 pages). See any current masthead page for ordering and Internet access instructions.

JA9621152